

日 本 国 特 許 庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office

出 願 年 月 日

Date of Application:

2000年 5月25日

出 願 番 号

Application Number:

特願2000-154045

出 願 人

Applicant(s):

株式会社日立製作所

CERTIFIED COPY OF
PRIORITY DOCUMENT

2001年 5月18日

特許庁長官
Commissioner,
Japan Patent Office

及川耕造



【書類名】 特許願

【整理番号】 K00003381

【提出日】 平成12年 5月25日

【あて先】 特許庁長官殿

【国際特許分類】 G06F 3/06

【発明者】

【住所又は居所】 神奈川県小田原市国府津 2 8 8 0 番地 株式会社日立製
作所 ストレージシステム事業部内

【氏名】 葛城 栄寿

【発明者】

【住所又は居所】 神奈川県小田原市国府津 2 8 8 0 番地 株式会社日立製
作所 ストレージシステム事業部内

【氏名】 中野 俊夫

【発明者】

【住所又は居所】 神奈川県小田原市国府津 2 8 8 0 番地 株式会社日立製
作所 ストレージシステム事業部内

【氏名】 岡見 吉規

【発明者】

【住所又は居所】 神奈川県小田原市国府津 2 8 8 0 番地 株式会社日立製
作所 ストレージシステム事業部内

【氏名】 佐藤 孝夫

【特許出願人】

【識別番号】 000005108

【氏名又は名称】 株式会社日立製作所

【代理人】

【識別番号】 100080001

【弁理士】

【氏名又は名称】 筒井 大和

【電話番号】 03-3366-0787

【手数料の表示】

【予納台帳番号】 006909

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【ブルーフの要否】 要

【書類名】 明細書

【発明の名称】 データ多重化方法およびデータ多重化システム

【特許請求の範囲】

【請求項1】 第1のホストコンピュータと第1の記憶装置により構成される第1の情報処理システムと、第2のホストコンピュータと第2の記憶装置により構成される少なくとも一つの第2の情報処理システムとをデータ転送バスを介して接続し、前記第1の情報処理システムで発生した第1の更新データを前記第2の情報処理システムにコピーすることで前記第1および第2の情報処理システムで同一のデータを多重に保持するデータ多重化方法であって、

前記第1の情報処理システムの稼働停止時に前記第1の情報処理システムの処理を引き継いで実行する前記第2の情報処理システムで発生する第2の更新データを識別する差分管理情報を前記第2の情報処理システムに持たせ、前記第1の情報処理システムの稼働再開後に、前記差分管理情報に基づいて、前記第2の更新データを選択的に前記第1の情報処理システムにコピーすることを特徴とするデータ多重化方法。

【請求項2】 第1のホストコンピュータと第1の記憶装置により構成される第1の情報処理システムと、第2のホストコンピュータと第2の記憶装置により構成される少なくとも一つの第2の情報処理システムをデータ転送バスを介して接続し、前記第1の情報処理システムで発生した第1の更新データを前記第2の情報処理システムに非同期にコピーするとともに、前記第2の情報処理システムに未コピーの前記第1の更新データを識別する第1の差分管理情報を前記第1の情報処理システムにて持つことで前記第1および第2の情報処理システムで同一のデータを多重に保持するデータ多重化方法であって、

前記第1の情報処理システムの稼働停止時に前記第1の情報処理システムの処理を引き継いで実行する前記第2の情報処理システムで発生する第2の更新データを識別する第2の差分管理情報を前記第2の情報処理システムに持たせ、前記第1の情報処理システムの稼働再開後に、前記第1および第2の差分管理情報にて特定される範囲のデータを選択的に前記第1の情報処理システムにコピーすることを特徴とするデータ多重化方法。

【請求項3】 第1のホストコンピュータと第1の記憶装置により構成される第1の情報処理システムと、第2のホストコンピュータと第2の記憶装置により構成される少なくとも一つの第2の情報処理システムをデータ転送バスを介して接続し、前記第1の情報処理システムで発生した第1の更新データを前記第2の情報処理システムに非同期にコピーするとともに、前記第2の情報処理システムに未コピーの前記第1の更新データを識別する第1の差分管理情報を前記第1の情報処理システムにて持つことで前記第1および第2の情報処理システムで同一のデータを多重に保持するデータ多重化方法であって、

前記第1の情報処理システムの稼働停止時に前記第1の情報処理システムの処理を引き継いで実行する前記第2の情報処理システムで発生する第2の更新データを識別する第2の差分管理情報を前記第2の情報処理システムに持たせ、前記第1の情報処理システムの稼働再開後に、前記第2の差分管理情報に基づいて、前記第2の更新データを選択的に前記第1の情報処理システムにコピーすることとを特徴とするデータ多重化方法。

【請求項4】 第1のホストコンピュータと第1の記憶装置により構成される第1の情報処理システムと、第2のホストコンピュータと第2の記憶装置により構成される少なくとも一つの第2の情報処理システムをデータ転送バスを介して接続し、前記第1の情報処理システムで発生した第1の更新データを前記第2の情報処理システムに同期してコピーすることで前記第1および第2の情報処理システムで常に同一のデータを多重に保持するデータ多重化方法であって、

前記第1の情報処理システムの稼働停止時に前記第1の情報処理システムの処理を引き継いで実行する前記第2の情報処理システムで発生する第2の更新データを識別する第2の差分管理情報を前記第2の情報処理システムに持たせ、前記第1の情報処理システムの稼働再開後に、前記第2の差分管理情報に基づいて、前記第2の更新データを選択的に前記第1の情報処理システムにコピーすることとを特徴とするデータ多重化方法。

【請求項5】 請求項1, 2, 3または4記載のデータ多重化方法において、前記差分管理情報または前記第1および第2の差分管理情報は、前記第1および第2の記憶装置の各々における複数のデータ格納単位の各々における前記第1

および第2の更新データのデータ多重化の完了の有無を示すビットマップであることを特徴とするデータ多重化方法。

【請求項6】 第1のホストコンピュータと第1の記憶装置により構成される第1の情報処理システムと、第2のホストコンピュータと第2の記憶装置により構成される少なくとも一つの第2の情報処理システムと、前記第1の情報処理システムと前記第2の情報処理システムとの間におけるデータ転送が行われるデータ転送バスとを含み、前記第1の情報処理システムにて発生した第1の更新データを、前記データ転送バスを経由して前記第2の情報処理システムへコピーすることで、前記第1の情報処理システムと前記第2の情報処理システムで同一のデータを多重に保持するデータ多重化システムであって、

前記第2の情報処理システムは、前記第1の情報処理システムが動作不能になった時に、前記第1の情報処理システムの処理を引き継いで実行する間に当該第2の情報処理システムで発生した第2の更新データを識別する差分管理情報と、前記第1の情報処理システムが動作可能になった時に、前記差分管理情報に基づいて前記第2の情報処理システムの前記第2の更新データを前記第1の情報処理システムに選択的にコピーする機能と、を備えたことを特徴とするデータ多重化システム。

【請求項7】 第1のホストコンピュータと第1の記憶装置により構成される第1の情報処理システムと、第2のホストコンピュータと第2の記憶装置により構成される少なくとも一つの第2の情報処理システムと、前記第1の情報処理システムと前記第2の情報処理システムとの間におけるデータ転送が行われるデータ転送バスとを含み、前記第1の情報処理システムにて発生した第1の更新データを、前記データ転送バスを経由して前記第2の情報処理システムへ非同期にコピーすることで、前記第1の情報処理システムと前記第2の情報処理システムで同一のデータを多重に保持するデータ多重化システムであって、

前記第1の情報処理システムは、前記第2の情報処理システムに未コピーの前記第1の更新データを識別する第1の差分管理情報を備え、

前記第2の情報処理システムは、前記第1の情報処理システムが動作不能になった時に、前記第1の情報処理システムの処理を引き継いで実行する間に当該第

2の情報処理システムで発生した第2の更新データを識別する第2の差分管理情報と、前記第1の情報処理システムが動作可能になった時に、前記第1の差分管理情報および前記第2の差分管理情報にて特定される範囲のデータを前記第1の情報処理システムに選択的にコピーする機能と、を備えたことを特徴とするデータ多重化システム。

【請求項8】 第1のホストコンピュータと第1の記憶装置により構成される第1の情報処理システムと、第2のホストコンピュータと第2の記憶装置により構成される少なくとも一つの第2の情報処理システムと、前記第1の情報処理システムと前記第2の情報処理システムとの間におけるデータ転送が行われるデータ転送バスとを含み、前記第1の情報処理システムにて発生した第1の更新データを、前記データ転送バスを経由して第2の情報処理システムへ非同期にコピーすることで、前記第1の情報処理システムと前記第2の情報処理システムで同一のデータを多重に保持するデータ多重化システムであって、

前記第1の情報処理システムは、前記第2の情報処理システムに未コピーの前記第1の更新データを識別する第1の差分管理情報を備え、

前記第2の情報処理システムは、前記第1の情報処理システムが動作不能になった時に、前記第1の情報処理システムの処理を引き継いで実行する間に当該第2の情報処理システムで発生した第2の更新データを識別する第2の差分管理情報と、前記第1の情報処理システムが動作可能になった時に、前記第2の差分管理情報に基づいて前記第2の情報処理システムの前記第2の更新データを前記第1の情報処理システムに選択的にコピーする機能と、を備えたことを特徴とするデータ多重化システム。

【請求項9】 第1のホストコンピュータと第1の記憶装置により構成される第1の情報処理システムと、第2のホストコンピュータと第2の記憶装置により構成される少なくとも一つの第2の情報処理システムと、前記第1の情報処理システムと前記第2の情報処理システムとの間におけるデータ転送が行われるデータ転送バスとを含み、前記第1の情報処理システムにて発生した第1の更新データを、前記データ転送バスを経由して第2の情報処理システムへ同期してコピーすることで、前記第1の情報処理システムと前記第2の情報処理システムで同

一のデータを多重に保持するデータ多重化システムであって、

前記第2の情報処理システムは、前記第1の情報処理システムが動作不能になった時に、前記第1の情報処理システムの処理を引き継いで実行する間に当該第2の情報処理システムで発生した第2の更新データを識別する第2の差分管理情報と、前記第1の情報処理システムが動作可能になった時に、前記第2の差分管理情報にて特定される範囲のデータを前記第1の情報処理システムに選択的にコピーする機能と、を備えたことを特徴とするデータ多重化システム。

【請求項10】 請求項6, 7, 8または9記載のデータ多重化システムにおいて、前記差分管理情報または前記第1および第2の差分管理情報は、前記第1および第2の記憶装置の各々における複数のデータ格納単位の各々における前記第1および第2の更新データのデータ多重化の完了の有無を示すビットマップであることを特徴とするデータ多重化システム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、データ多重化技術に関し、特に、非同期あるいは同期式のデータコピー技術を用いて、互いに独立な複数の情報処理システム間でデータを多重に保持する技術等に適用して有効な技術に関する。

【0002】

【従来の技術】

たとえば、金融機関、保険会社のように、大容量のデータを保持しかつデータの損失が業務遂行や社会生活に重大な影響を与えるユーザーにおいては、データのバックアップを取る事は従来から行われているが、近年、天災や事故の教訓から、遠隔地に配置したバックアップシステムにデータを保存する事が要望されるようになった。

【0003】

すなわち、マスターシステム上の更新データを、遠隔地に配置されたバックアップシステムへコピーし、マスターシステムが災害や障害等により運転を停止した際にはバックアップシステムに運転を切り替えて継続動作可能にすることで、

データの保全および稼働継続の保証を実現するものである。

【0004】

このような遠隔地間でのデータ多重化によるデータ保全等に関しては、例えば、特表平8-509565号公報に開示されている技術がある。その方法は、マスタシステムのデータをそのままバックアップシステムに反映することにより「ミラー状態」を維持するものであり、マスタシステムでの運用が不可能となった場合に、バックアップシステムに運用を移行し、システム運用再開を一段と容易にしようとするものである。バックアップシステムへのバックアップコピー方法としては、マスタシステムとバックアップシステム間のデータの更新契機から、大きく「同期型」と「非同期型」の2種類に分けることができる。

【0005】

同期型は、マスタシステムのホストより発生した各データ更新要求に対し、まずマスタシステムの記憶装置に書き込みを行い、続いてバックアップシステムの記憶装置に向けて書き込みを実施し、バックアップシステムより書き込み終了通知を受領したところで、マスタシステムの記憶装置はマスタシステムのホストに対して、最終的な書き込み終了報告を行い、常にマスタシステムとバックアップシステムのデータ更新は同期を保つ技術である。

【0006】

それに対し、非同期型は、マスタシステムのホストより発生した各データ更新要求に対し、マスタシステムの記憶装置に書き込みを終了した時点で、マスタシステムのホストに書き込み終了報告を行い、このデータ更新に対するバックアップシステムへのデータ更新の実行は遅れて、即ち、非同期に実施される技術である。

【0007】

非同期コピー方式の場合、マスタシステムに更新されたデータをバックアップシステムへ反映するまでの間、一時的にデータの差異が生じる。この為、バックアップシステムへのデータの未反映状態を管理する必要がある。例えば、特開平10-198607号公報に開示された方法は、論理トラック毎の差分の有無を示す差分管理テーブルを持ち、それを元に、バックアップシステムへ更新データ

をコピーするものである。

【0008】

通常「ミラー状態」を維持している状態において、仮にマスタシステムが重大災害等で運用不可能に陥った場合、「同期型」においては、データ更新の同期性から「ミラー状態」は維持される。この場合、最終更新に関するデータ分は、タイミングによりリモートシステムに反映されないかもしれないが、整合性に関しては問題ない。ところが、「非同期型」においては、バックアップシステムには非同期に更新データが反映されるため、マスタシステムとバックアップシステムに不整合が生じる。この不整合状態を回復する為、マスターシステムが災害や障害による停止状態から復旧した時に、バックアップシステムのデータの全てをマスターシステムへとコピーしなければならないが、大容量のデータのコピーには多くの時間を要し、結果的にマスターシステムの復旧までの時間を遅延させることになる。

【0009】

この不整合状態の回復方法としては、例えば、特開平6-290125号公報に開示されている方式がある。これは、更新データを待ち行列化して、更新があった順にバックアップシステムにコピーし、整合性を維持するようにしたものであり、障害から回復時する際には、バックアップシステム内の保留書き込み待ち行列により、不整合状態を回復するものである。

【0010】

また、米国特許第5857208号には、ローカルシステムとリモートシステムの間でデータ多重化を行うシステムにおいて、リモートシステムに磁気テープ装置等のバックアップユニットを接続し、特定時点にて、ローカルシステムからリモートシステムへのデータコピーを停止してバックアップユニットにデータを複写する際に発生するローカルシステムとリモートシステムとの間の差分データを、ローカルシステム側に設けられたトラックステータスにて管理し、リモートシステムのバックアップ完了後に、差分データをローカルシステムからリモートシステムにコピーして整合性を回復する技術が示されている。

【0011】

【発明が解決しようとする課題】

マスターシステムが災害や障害により停止した際には、マスターシステムが回復するまでの間はバックアップシステムに切り替え業務を継続する必要がある。非同期コピーの場合、前述の様にバックアップシステムへ切り替えた時点で、マスターシステムとの差分（これを第1の差分と呼ぶことにする）が生じているが、バックアップシステムで業務を続けていると、今度はマスターシステムへの未反映データである、第2の差分が生じる。この第2の差分の量は、マスターシステムが回復するまでの時間によって異なり、また第1と第2の差分は、部分的あるいは全部が一致しているかもしれない。この不整合状態を回復する為、マスターシステムが復旧した時に、バックアップシステムのデータの全てをマスターシステムへとコピーする方法では、多くの時間を要し、結果的にマスターシステムの復旧までの時間を遅延させることになる。

【0012】

本発明の目的は、マスターシステムが回復した際に、速やかにマスターシステムおよびバックアップシステムの各々におけるデータの不整合を解消し、マスターシステムの運転再開までの時間を短縮することにある。

【0013】

本発明の他の目的は、複数の情報処理システム間で同一のデータを多重に保持することでデータ保証および稼働継続保証を実現するデータ多重化システムにおいて、一部の情報処理システムの稼働停止および稼働再開に伴う各情報処理システム間でのデータの不整合の回復所要時間を短縮することにある。

【0014】

本発明の他の目的は、複数の情報処理システム間で同一のデータを多重に保持することでデータ保証および稼働継続保証を実現するデータ多重化システムにおいて、一部の情報処理システムの稼働停止から稼働再開までの所要時間を短縮することにある。

【0015】

【課題を解決するための手段】

本発明は、第1のホストコンピュータと第1の記憶装置により構成される第1

の情報処理システムと、第2のホストコンピュータと第2の記憶装置により構成される少なくとも一つの第2の情報処理システムとをデータ転送バスを介して接続し、第1の情報処理システムで発生した第1の更新データを第2の情報処理システムにコピーすることで第1および第2の情報処理システムで同一のデータを多重に保持するデータ多重化方法において、第1の情報処理システムの稼働停止時に第1の情報処理システムの処理を引き継いで実行する第2の情報処理システムで発生する第2の更新データを識別する差分管理情報を第2の情報処理システムに持たせ、第1の情報処理システムの稼働再開後に、差分管理情報に基づいて、第2の更新データを選択的に第1の情報処理システムにコピーするものである。

【0016】

また、本発明は、第1のホストコンピュータと第1の記憶装置により構成される第1の情報処理システムと、第2のホストコンピュータと第2の記憶装置により構成される少なくとも一つの第2の情報処理システムをデータ転送バスを介して接続し、第1の情報処理システムで発生した第1の更新データを第2の情報処理システムに非同期にコピーするとともに、第2の情報処理システムに未コピーの第1の更新データを識別する第1の差分管理情報を第1の情報処理システムにて持つことで第1および第2の情報処理システムで同一のデータを多重に保持するデータ多重化方法において、第1の情報処理システムの稼働停止時に第1の情報処理システムの処理を引き継いで実行する第2の情報処理システムで発生する第2の更新データを識別する第2の差分管理情報を第2の情報処理システムに持たせ、第1の情報処理システムの稼働再開後に、第1および第2の差分管理情報にて特定される範囲のデータを選択的に第1の情報処理システムにコピーするものである。

【0017】

また、本発明は、第1のホストコンピュータと第1の記憶装置により構成される第1の情報処理システムと、第2のホストコンピュータと第2の記憶装置により構成される少なくとも一つの第2の情報処理システムと、第1の情報処理システムと第2の情報処理システムとの間におけるデータ転送が行われるデータ転送

バスとを含み、第1の情報処理システムにて発生した第1の更新データを、データ転送バスを経由して第2の情報処理システムへコピーすることで、第1の情報処理システムと第2の情報処理システムで同一のデータを多重に保持するデータ多重化システムにおいて、第2の情報処理システムは、第1の情報処理システムが動作不能になった時に、第1の情報処理システムの処理を引き継いで実行する間に当該第2の情報処理システムで発生した第2の更新データを識別する差分管理情報と、第1の情報処理システムが動作可能になった時に、差分管理情報に基づいて第2の情報処理システムの第2の更新データを第1の情報処理システムに選択的にコピーする機能と、を備えたものである。

【0018】

また、本発明は、第1のホストコンピュータと第1の記憶装置により構成される第1の情報処理システムと、第2のホストコンピュータと第2の記憶装置により構成される少なくとも一つの第2の情報処理システムと、第1の情報処理システムと第2の情報処理システムとの間におけるデータ転送が行われるデータ転送バスとを含み、第1の情報処理システムにて発生した第1の更新データを、データ転送バスを経由して第2の情報処理システムへ非同期にコピーすることで、第1の情報処理システムと第2の情報処理システムで同一のデータを多重に保持するデータ多重化システムにおいて、第1の情報処理システムは、第2の情報処理システムに未コピーの第1の更新データを識別する第1の差分管理情報を備え、第2の情報処理システムは、第1の情報処理システムが動作不能になった時に、第1の情報処理システムの処理を引き継いで実行する間に当該第2の情報処理システムで発生した第2の更新データを識別する第2の差分管理情報と、第1の情報処理システムが動作可能になった時に、第1の差分管理情報および第2の差分管理情報にて特定される範囲のデータを第1の情報処理システムに選択的にコピーする機能と、を備えたものである。

【0019】

より具体的には、一例として、ホストコンピュータと記憶装置により構成される第1のシステムと、ホストコンピュータと記憶装置により構成される複数の第2のシステムと、第1のシステムと第2のシステムの間のデータ転送を行うデー

タ転送バスから成り、第1のシステムに対する更新データを、第2のシステムへと非同期あるいは同期してコピーしデータの多重化を行い、第1のシステムが動作不能となった時に、第2のシステムが処理を継続できるデータ多重化システムにおいて、

第1のシステムは、必要に応じて、第1のシステムに対する更新データを、第2のシステムへコピーする際に生じる差分を第1の差分情報として管理する機能を有しており、

第2のシステムは、第1のシステムが動作不能状態となってから回復に至るまでに発生する、第2のシステムに対する更新データを、第2の差分データとして管理し、第1のシステムが動作可能となった時に、前記第1の差分情報と前記第2の差分情報の範囲あるいは、前記第2の差分情報の範囲のみを、非同期に第1のシステムへとコピーする機能を有する。

【 0 0 2 0 】

【発明の実施の形態】

以下、本発明の実施の形態を図面を参照しながら詳細に説明する。

【 0 0 2 1 】

（実施の形態1）

図1は本発明の一実施の形態であるデータ多重化方法を実施するデータ多重化システムの一例である情報処理システムの構成例を示すブロック図である。

【 0 0 2 2 】

図1に例示されるように、本実施の形態の情報処理システムは、マスターシステム100と、バックアップシステム400の2つのシステムから構成されている。

【 0 0 2 3 】

マスターシステム100は、ホストコンピュータ110とディスク制御装置120とディスク記憶装置230から構成されている。ディスク制御装置120は、チャンネルI/F140、141、チャンネル処理プロセッサ150、キャッシュメモリ160、共用メモリ170、内部バス240、ドライブ処理プロセッサ200、ドライブI/F210を備えており、ホストコンピュータ110とは、チ

ヤネルバス130を通して結合されている。また、ディスク記憶装置230はドライブ転送バス250を通して、ドライブI/F210と結合されている。ディスク制御装置120内のチャンネル処理プロセッサ150は、ホストコンピュータ110とのデータ送受信の際と、ホストコンピュータ110からの更新データをチャンネルI/F141を用いてバックアップシステム400へと転送する際に機能する。ディスク制御装置120内のドライブ処理プロセッサ200は、ドライブI/F210を通して、ディスク記憶装置230とのデータ送受信の際に機能する。

【0024】

共用メモリ170内の差分管理テーブル180は、チャンネル処理プロセッサ150によって更新・参照され、バックアップシステム400への未反映データを管理する為に用いられる。また、共用メモリ170内のペア状態管理テーブル190はコピー状態を管理する為に用いられる。

【0025】

すなわち、本実施の形態のマスターシステム100のディスク制御装置120は、ホストコンピュータ110から更新データを受け取ると、当該ディスク制御装置120内のキャッシュメモリ160への当該更新データの書き込み処理が完了した時点で、ホストコンピュータ110に対して書き込み完了を報告し、チャンネルI/F141、インターフェイスケابل600を介したバックアップシステム400への当該更新データのコピー処理は後の任意契機で行う非同期コピー処理が行われる。

【0026】

バックアップシステム400は、マスターシステム100と全く同様の構成・機能を有しており、ホストコンピュータ410と、その配下で稼働するディスク制御装置420およびディスク記憶装置530からなる。バックアップシステム400では、インターフェイスケابل600に接続されるチャンネルI/F440を介してマスターシステム100との間のデータの授受が行われ、チャンネルI/F441を介してホストコンピュータ410との間のデータの授受が行われる。

【0027】

本実施の形態の場合、ディスク制御装置420において、共用メモリ470内には、差分管理テーブル480が設けられている。この差分管理テーブル480は、チャンネル処理プロセッサ450によって内部バス540を介して更新・参照されるものであり、マスターシステム100が障害等により運転停止中に、ホストコンピュータ410からの更新データにより発生する、差分データを管理するものである。

【0028】

マスターシステム100とバックアップシステム400との間のデータ転送はインターフェイスケابل600を通して行われ、ホストコンピュータ110とホストコンピュータ410間の問い合わせは、ホスト間通信ケーブル700により行われる。なお、このインターフェイスケابل600やホスト間通信ケーブル700は、バックアップシステム400を遠隔地に配置する場合には、両者間を接続する光ファイバや通信回線、さらには無線通信の電波等の情報通信媒体が考えられる。

【0029】

図2は、本実施の形態の情報処理システムにて用いられる差分管理テーブル180および差分管理テーブル480の構成の一例を示す概念図である。本実施の形態においては、ホストコンピュータ110および410からディスク制御装置120および420へのI/Oは、論理デバイス番号(DEV No.)、論理シリンダ番号(CYL No.)、論理ヘッド番号(HD No.)によりアドレスングされ、論理デバイス番号(DEV No.)は0からmの範囲を、論理シリンダ番号(CYL No.)は0からnの範囲を、論理ヘッド番号(論理トラック番号)(HD No.)は0から14の値を取りうる。

【0030】

図2に例示される差分管理テーブル180および480は、一例としてビットマップにて構成され、論理ヘッド番号は0から14を1ワード(16ビット)の情報として管理する(右端のビットは管理を容易にする為未使用としている)。この1ワード情報をn+1個で1つの論理デバイスの差分状態を表し、それがm

+1個で全デバイス数分の差分状態を表す。したがって、図2のテーブルによれば、論理ヘッド番号（論理トラック番号）（HD No.）を最小単位として、差分有りまたは差分無しという事を管理出来る。

【0031】

図3から図6は本実施の形態の情報処理システムの作用の一例を示すフローチャートである。以下フローチャートに基づき、処理の流れを説明する。

【0032】

図3は、マスターシステム100が処理不能となった時のバックアップシステム400が縮退運転を開始する迄の処理の流れを示す。まず、ステップ1000で、ホストコンピュータ410はマスターシステム100が動作不能状態となった事を検出する。これは、例えばホスト間通信ケーブル700を用い、ホストコンピュータ410が定期的に、ホストコンピュータ110と通信し、状態チェックする事で実現される。次にステップ1010で、ホストコンピュータ410は、ディスク制御装置420に対して、SWAP要求を発行する。このSWAP要求とは、通常バックアップとしての動作を行っているディスク制御装置420に、マスターシステム100内のディスク制御装置120と同等の動作を行わせるものである。このSWAP要求を受けたディスク制御装置420は、ステップ1020で、差分管理テーブル480をクリアし、ステップ1030で、ベア状態管理テーブル490にベア状態としてSWAPを格納する。ベア状態：SWAPとはバックアップシステム400が縮退運転を開始している状態で、この状態の時には、ディスク制御装置420は、ホストコンピュータ410からの更新データは全て差分データとして管理し、マスターシステム100へのコピーは行わない状態である。

【0033】

このSWAP状態の際の、ディスク制御装置420の処理の流れを図4のフローチャートを用いて説明する。まず、ステップ2000でホストコンピュータ410からコマンドを受領し、ステップ2010でそのコマンドがライトコマンドでない場合はステップ2030で対応するコマンド処理を実行する。ライトコマンドの場合はステップ2040で、ベア状態管理テーブル490を参照し、ベア

状態がSWAPあるいは、後述のCOPYの時には、2050に進む。状態がCOPYでない場合はステップ2070に進む。ステップ2050では、ライト対象のアドレスから、対応する差分情報のアドレスを算出し、ステップ2060で差分管理テーブル480の当該ビットをON（“1”）する。次にステップ2070で、当該ライトデータをキャッシュメモリ460に格納し、ステップ2080で、ホストコンピュータ410にライト完了を報告し、ライトコマンドの処理を継続させる。なお、キャッシュメモリ460に格納したライトデータは、ドライブ処理プロセッサ500により非同期にディスク記憶装置530に格納する。

【0034】

次に、マスターシステム100が運転不能な状態から回復した場合の、バックアップシステム400およびマスターシステム100の動作を図5、図6のフローチャートを用いて説明する。

【0035】

まず、図5のステップ3000で、ホストコンピュータ410はマスターシステム100が回復した事を検出する。これは、例えばホスト間通信ケーブル700を用い、ホストコンピュータ410が定期的に、ホストコンピュータ110の状態をチェックする事で実現される。

【0036】

次にステップ3010で、ホストコンピュータ410から、ディスク制御装置420に対して、Resync要求を発行する。Resync要求を受け付けたディスク制御装置420は、ステップ3020で、インターフェイスケابل600を通して、ディスク制御装置120内の差分管理テーブル180のデータを、ディスク制御装置420内にリードし、差分管理テーブル480との論理和を演算した結果を差分管理テーブル480に格納する。この新たに生成された差分情報が、マスターシステム100とバックアップシステム400の整合性を維持するために必要となる。

【0037】

次にステップ3030で、ベア状態管理テーブル490内のベア状態をCOP

Yに変更する。次にステップ3040で差分コピー処理を起動する。差分コピー処理は、ホストコマンドからの更新データとは非同期に実行されるJOBであり、ステップ3040で、差分コピーJOBを起動しても良いし、差分コピーJOB自体が、ベア状態管理テーブル490を定期的に参照して、ベア状態がCOPYとなったのを契機に、コピーを開始してもよい。

【0038】

差分コピーの処理は、まずステップ3050で、差分管理テーブル480を参照し、ビットONとなっている個所を検索する。次にステップ3060で、ビットONの個所が見つかったら、ステップ3070に進み、当該差分データをディスク制御装置120に転送し、転送済の当該ビットをOFF（“0”）にする。ステップ3060で、ビットONの個所が見つからなかったら、ステップ3050に進み、再び差分のサーチを実行する。

【0039】

図6のフローチャートは、バックアップシステム400での処理を、再びマスターシステム100へと移行するシーケンスの一例である。まず、ステップ4000で、ホストコンピュータ410はI/Oを停止する。これは移行シーケンス中に新たな差分が出来ると、その部分に関しては整合性が保証出来ないからである。次にステップ4010で、ホストコンピュータ410より、ディスク制御装置420に対して、DELPAIR要求を発行する。

【0040】

ディスク制御装置420は、DELPAIR要求を受信した時、ステップ4020で、差分管理テーブル480を用いて差分をサーチし、ステップ4030で、差分データが検出された時には、ステップ4040で当該差分データをディスク制御装置120へ転送する。差分データが全く無くなった時には、ステップ4050に進み、ベア状態管理テーブル490内のベア状態を「通常」に設定する。次にステップ4060で、ホストコンピュータ410に対し、DELPAIR要求の完了を報告する。

【0041】

DELPAIR要求の完了報告を受け、ステップ4070でホストコンピュー

タ410からホストコンピュータ110へ制御を移行する。次にステップ4080で、ホストコンピュータ110からディスク制御装置120に対し、ベア回復要求を発行する。ディスク制御装置120は、ベア回復要求を受け、ステップ4090でディスク制御装置120内の差分管理テーブル180を全てクリアし、内部状態をCOPYに変更する。これにより、前述の図5のステップ3050から3060と同様に差分コピーの処理を開始する。この時点でマスターシステム100とバックアップシステム400の格納データの整合性は保たれているから、通常、ベアを確立する為に必要な全データコピーは必要ない。以上の処理により、マスターシステム100が運転を停止する前の状態に回復される。

【0042】

本実施の形態のデータ多重化方法およびデータ多重化システムによれば、マスターシステム100が障害等による運転停止状態から運転可能状態に回復した際に、バックアップシステム400からマスターシステム100への必要最小限のデータコピーにより、速やかにマスターシステム100とバックアップシステム400におけるデータを整合状態とする事が可能となり、マスターシステム100の運転再開までの時間を短縮することが可能になる。

【0043】

(実施の形態2)

本発明の第2の実施の形態として、マスターシステム100の復旧に際して、バックアップシステム400からマスターシステム100へのコピー対象となる差分データを生成する際に、マスターシステム100の差分管理テーブル180の差分情報はマージしない場合を示す。

【0044】

すなわち、この第3の実施の形態においては、図5のステップ3020の処理は不要である。それ以外の部分については、第1の実施形態と同様である。

【0045】

これは、業務によっては、マスターシステム100の稼働停止までの間に発生した、当該マスターシステム100とバックアップシステム400との間における更新データの不整合（未反映データの存在）を許容する場合も考えられるから

である。

【0046】

この場合には、上述の第1の実施の形態と同様の効果が得られるとともに、マスターシステム100側の差分管理テーブル180に基づく差分データの復旧は行われないので、バックアップシステム400からマスターシステム100への差分データのコピー処理の所要時間は、第1の実施の形態の場合よりも短縮される。

【0047】

(実施の形態3)

また、図1等を参照して、本発明の第3の実施の形態の情報処理システムとして、マスターシステム100からバックアップシステム400に同期コピーを行なう場合を説明する。

【0048】

すなわち、この第3の実施の形態における同期コピーの場合には、マスターシステム100のディスク制御装置120は、ホストコンピュータ110から更新データを受け取ると、当該ディスク制御装置120内のキャッシュメモリ160（ディスク記憶装置230）への当該更新データの書き込み処理を行うとともに、チャンネルI/F141、インターフェイスケابل600を介したバックアップシステム400への当該更新データのコピー処理を実行し、マスターシステム100およびバックアップシステム400の両方において当該更新データの書き込み完了が確認された後に、要求元のホストコンピュータ110に対して書き込み完了を報告する。

【0049】

従って、この第2の実施の形態においては、図1のマスターシステム100側の差分管理テーブル180は不要である。また図5のステップ3020の処理は不要である。それ以外の部分については、第1の実施形態と同様である。

【0050】

この第3の実施の形態の場合にも、上述の第1の実施の形態と同様の効果が得られるとともに、マスターシステム100側の差分管理テーブル180は存在せ

ず、当該差分管理テーブル 180 による差分データの復旧は存在しないので、バックアップシステム 400 からマスターシステム 100 への差分データのコピー処理の所要時間は、第 1 の実施の形態の場合よりも短縮される。

【0051】

以上本発明者によってなされた発明を実施の形態に基づき具体的に説明したが、本発明は前記実施の形態に限定されるものではなく、その要旨を逸脱しない範囲で種々変更可能であることはいうまでもない。

【0052】

たとえば、上述の実施の形態では、マスターシステムとバックアップシステムが 1 対 1 の場合を例示したが、バックアップシステムが複数存在してもよい。

【0053】

【発明の効果】

本発明のデータ多重化技術によれば、マスターシステムが回復した際に、速やかにマスターシステムおよびバックアップシステムの各々におけるデータの不整合を解消し、マスターシステムの運転再開までの時間を短縮することができる、という効果が得られる。

【0054】

本発明のデータ多重化技術によれば、複数の情報処理システム間で同一のデータを多重に保持することでデータ保証および稼働継続保証を実現するデータ多重化システムにおいて、一部の情報処理システムの稼働停止および稼働再開に伴う各情報処理システム間でのデータの不整合の回復所要時間を短縮することができる、という効果が得られる。

【0055】

本発明のデータ多重化技術によれば、複数の情報処理システム間で同一のデータを多重に保持することでデータ保証および稼働継続保証を実現するデータ多重化システムにおいて、一部の情報処理システムの稼働停止から稼働再開までの所要時間を短縮することができる、という効果が得られる。

【図面の簡単な説明】

【図 1】

本発明の一実施の形態であるデータ多重化方法を実施するデータ多重化システムの一例である情報処理システムの構成例を示すブロック図である。

【図 2】

本発明の一実施の形態である情報処理システムにて用いられる差分管理テーブルの構成の一例を示す概念図である。

【図 3】

本発明の一実施の形態である情報処理システムの作用の一例を示すフローチャートである。

【図 4】

本発明の一実施の形態である情報処理システムの作用の一例を示すフローチャートである。

【図 5】

本発明の一実施の形態である情報処理システムの作用の一例を示すフローチャートである。

【図 6】

本発明の一実施の形態である情報処理システムの作用の一例を示すフローチャートである。

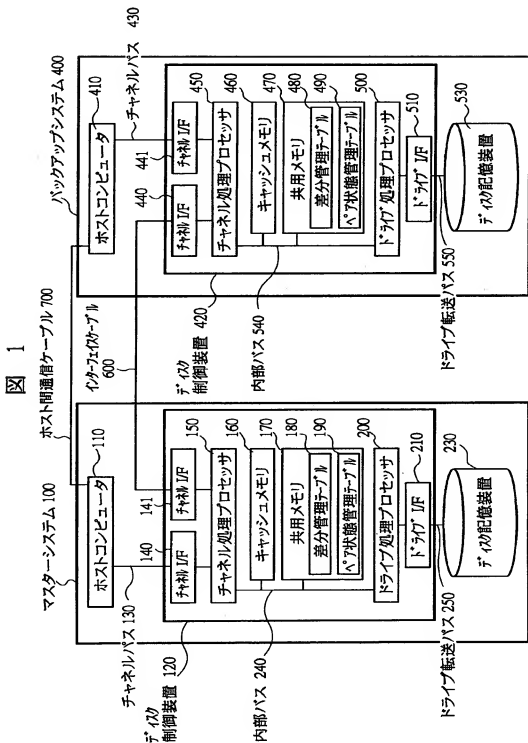
【符号の説明】

1 0 0 … マスターシステム（第 1 の情報処理システム）、1 1 0 … ホストコンピュータ（第 1 のホストコンピュータ）、1 2 0 … ディスク制御装置（第 1 の記憶装置）、1 3 0 … チャネルバス、1 4 0 … チャネル I / F、1 4 1 … チャネル I / F、1 5 0 … チャネル処理プロセッサ、1 6 0 … キャッシュメモリ、1 7 0 … 共用メモリ、1 8 0 … 差分管理テーブル（第 1 の差分管理情報）、1 9 0 … ベア状態管理テーブル、2 0 0 … ドライブ処理プロセッサ、2 1 0 … ドライブ I / F、2 3 0 … ディスク記憶装置（第 1 の記憶装置）、2 4 0 … 内部バス、2 5 0 … ドライブ転送バス、4 0 0 … バックアップシステム（第 2 の情報処理システム）、4 1 0 … ホストコンピュータ（第 2 のホストコンピュータ）、4 2 0 … ディスク制御装置（第 2 の記憶装置）、4 4 0 … チャネル I / F、4 4 1 … チャネル I / F、4 5 0 … チャネル処理プロセッサ、4 6 0 … キャッシュメモリ、4 7 0

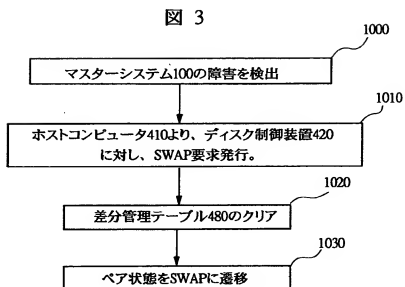
…共用メモリ、480…差分管理テーブル（第2の差分管理情報）、490…ベ
ア状態管理テーブル、500…ドライブ処理プロセッサ、530…ディスク記憶
装置（第2の記憶装置）、540…内部バス、600…インターフェイスケーブ
ル、700…ホスト間通信ケーブル。

【書類名】 図面

【図 1】

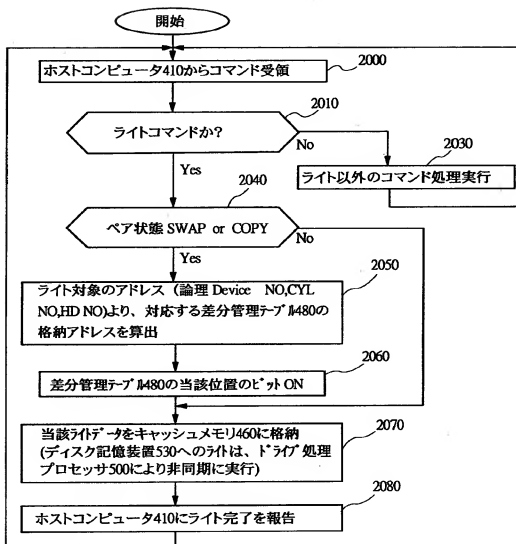


【図 3】



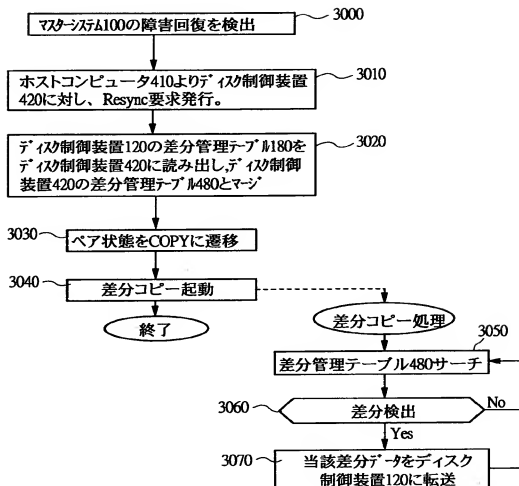
【図 4】

図 4



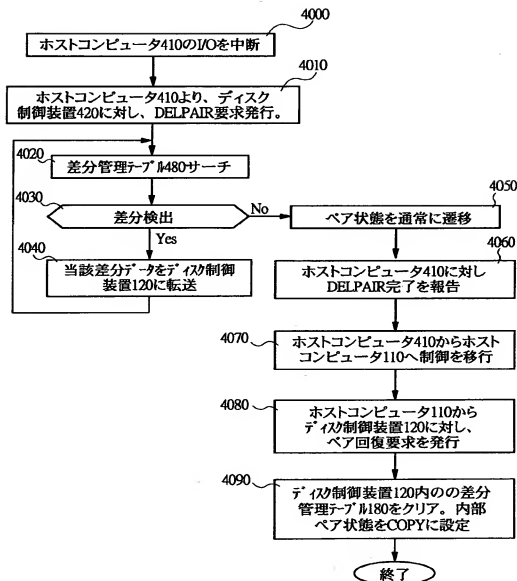
【図5】

図 5



【図 6】

図 6



【書類名】 要約書

【要約】

【課題】 マスターシステムが回復した際に、速やかにマスターシステムおよびバックアップシステムの各々におけるデータの不整合を解消し、マスターシステムの運転再開までの時間を短縮する。

【解決手段】 マスターシステム100とバックアップシステム400からなるデータ多重化システムにおいて、バックアップシステム400は、マスターシステム100が動作不能状態となってから回復に至るまでの代行稼働期間に発生する、バックアップシステム400での更新データを、差分管理テーブル480にて第2の差分情報として管理し、マスターシステム100が動作可能となった時に、マスターシステム100内の差分管理テーブル180の第1の差分情報と、差分管理テーブル480の第2の差分情報を併せた範囲あるいは、第2の差分情報の範囲のみを、マスターシステム100にコピーしてデータの不整合を解消する。

【選択図】 図1

出 願 人 履 歷 情 報

識別番号 [000005108]

1. 変更年月日 1990年 8月31日

[変更理由] 新規登録

住 所 東京都千代田区神田駿河台4丁目6番地

氏 名 株式会社日立製作所